REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
17-08-1999	Technical Paper	

4. TITLE AND SUBTITLE

Classification of Organophosphorus Compound Infrared Spectra by Pattern Recognition Techniques

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Dr. Howard T. Mayfield, DeLyle Eastwood and Larry A. Burggraf

5d. PROJECT NUMBER

OAFT0062

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory (MLQL)
139 Barnes Drive, Suite 2
Tyndall AFB FL 32403

8. PERFORMING ORGANIZATION REPORT NUMBER

AFRL-ML-TY-TP-2003-4513

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

N/A

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

N/A

12. DISTRIBUTION/AVAILABILITY STATEMENT

Distribution Unlimited

13. SUPPLEMENTARY NOTES

N/A

14. ABSTRACT

Infrared spectra of organophosphorus compounds, including pesticides and a set of neurotoxins which have been banned from use by international agreement, along with their precursors and hydrolysis products, were obtained from a variety of sources. The data were treated to minimize spectral information related to the spectral orgin. A common spectral wavelength range was selected and spectral data within this range wre transduced into data vectors. Computer-assisted classification tools wre used to classify the spectra as pesticides versus neurotoxins and their precursors and hydrolysis products. The performance of a k-nearest neighbor classifier for this distinction is compared with several artificial neural network classifiers.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Andrew Poulis
Unclassified	Unclassified	Unclassified	UL	19	19b. TELEPHONE NUMBER *(Include area code)* (850) 283-6285

Classification of Organophosphorus Compound Infrared Spectra by Pattern Recognition Techniques

Howard T. Mayfield
Air Force Research Laboratory, Airbase and Environmental Technology Division, AFRL/MLQR, 139 Barnes Drive, Suite 2, Tyndall AFB, FL 32403-5323

DeLyle Eastwood, and Larry A. Burggraf
Air Force Institute of Technology, Department of Engineering Physics, 2950 P. Street, Wright-Patterson AFB, OH, 45433.

Abstract:

Infrared spectra of organophosphorus compounds, including pesticides and a set of neurotoxins which have been banned from use by international agreement, along with their precursors and hydrolysis products, were obtained from a variety of sources. The data were treated to minimize spectral information related to the spectral origin. A common spectral wavelength range was selected and spectral data within this range were transduced into data vectors. Computer-assisted classification tools were used to classify the spectra as pesticides versus neurotoxins and their precursors and hydrolysis products. The performance of a k-nearest neighbor classifier for this distinction is compared with several artificial neural network classifiers.

Introduction:

There is an interest in analyzing and differentiating among different classes of organophosphorus compounds for environmental monitoring. Organophosphorus compounds differ greatly in toxicity, ranging from relatively nontoxic insecticides and herbicides to more toxic pesticides to extremely lethal neurotoxins. One reason for the increased use of organophosphorus pesticides relative to other pesticides is that they are less persistent and hydrolyze fairly quickly.[1]

Organophosphorus compounds include all organic compounds involving the heteroatom phosphorus. An important class of these compounds is made up of esters of phosphoric acid, H_3PO_4, and related phosphorus containing acids. A wide variety of structures are possible, some of which are shown in Figure 1. Several of the phosphate esters are acetylcholinesterase inhibitors, which acetylcholinesterase inhibitors, which operate as neurotoxins by disrupting the transmission of nerve

impulses. A small subset of the phosphate esters so efficiently inhibits acetylcholinesterase as to be deadly to higher mammals. Because of their danger to humans, certain organophosphorus compounds have now been banned from production by international agreement. The ban includes certain organophosphorus compounds which are neurotoxic to humans as well as their precursors and hydrolysis products.[2] Most organophosphorus pesticides belong to the structural classes of organophosphates, organophosphonates, and organthiophosphonates.[1] Figure 2 shows the structures of four neurotoxins which are now banned under international convention: dimethylphosphoramidocyanidic acid; ethyl ester, methylphosphonofluoridic acid, (1-methylethyl) ester; methylphosphonofluoridic acid, 1,2,2-trimethylpropyl ester; and methylphosphonothioic acid, S-[2-[bis(1-methylethyl)amino]ethyl] O-ethyl ester.[1-4]

Infrared (IR) spectroscopy, because of its selectivity for chemical structure, is ideal for distinguishing organophosphorus compounds using pattern recognition. Infrared spectrometers can be made compact and portable to support on-site detection and identification efforts for field, remote or in situ measurements.[5,6]

Pattern Recognition Techniques:

Pattern recognition techniques are chemometric computational techniques used to assign spectra into distinct classes depending upon multivariate measurements. To apply pattern recognition techniques, a set of spectra must be available which are already known to belong to specific classes. Infrared spectra data vectors are created by assigning absorbance data to elements of the data vector in order by frequency or wavelength.[7-9] Most pattern recognition algorithms require that the number of training set data vectors greatly exceed the dimensionality used in the final classification. It is usually necessary to either select specific measurements from the spectra or to combine data from adjacent measurements into a single measurement. A common approach to combining adjacent spectral data is to employ a spectral bin that contains the sum, the average, or the maximum spectral value within a certain spectral range. Bins may vary in size or they may be held to a given bandwidth throughout the spectrum. We used average absorbance assigned to constant width frequency bins.

Artificial Neural Networks (ANNs) are computational constructs in which multivariate inputs are used to derive new multivariate layered structures which are used in turn to formulate multivariate outputs. The structure is composed of nodes (or neurons) and connections between them in a layered structure. Nodes are connected to nodes in other layers by means of mathematical combinations.If the ANN has been fully implemented and has been properly trained, the answers are obtained at the nodes of the output layer.[10] For classification purposes, it is common to assign one output layer node to each class, and denote an input pattern's membership in a class by a value of 1.0 in the corresponding output node.[11] There may be one or more hidden layers summarized by an intermediate layer.

A typical hidden layer node can be formulated as shown in Equation 1, where m is the number of input nodes, N_j is the value of node j, x_i is the value of input node i, w_i is the value of corresponding weight, b_i is the bias of the node, and f_j is known as the transfer function. Typical transfer functions used in classification problems include the log-sigmoid function, as is shown in Equation 2, the tan-sigmoid function, shown in Equation 3, or linear functions. The log-sigmoid function varies from 0 to 1 over a range of $-\infty$ to $+\infty$. The tan sigmoid function varies from -1 to $+1$ over the range of $-\infty$ to $+\infty$. Nodes of subsequent layers can be similarly formatted except x_i values become the values of the nodes of the preceding layer. [10,12]

$$N_j = f_j\left(\sum_{i=1}^{m} w_i x_i + b_i\right) \tag{1}$$

$$f(x) = \frac{1}{1+e^{-x}} \tag{2}$$

$$f(x) = \frac{2}{1-e^{-2x}} - 1 \tag{3}$$

Much of the challenge in implementing neural networks lies in calculating the weights and biases of the combinations so that the network output is correct. In the classical back-propagation approach, error signals from the output layer during training are propagated through the weights and biases of the preceding layers in a series of iterations until the network yields correct classifications.[10,12]

Several training strategies have been developed in effort to optimize the training process in terms of time and memory requirements. Typically, weights and biases are randomly initialized and are then adjusted in an iterative process in which the network is evaluated using known input data and errors are obtained and fed back into the network from output layer toward the input layer in a process termed backpropagation. A number of algorithms are available for performing the backpropagation process, and they represent tradeoffs in terms of memory requirements, execution speed, and ability to find global optimum conditions versus local optima. The gradient descent backpropagation training system, developed by Rumelhart, *et al.*,[13] has been widely used in previous neural network-classification experiments with chemical data. This system utilizes a steepest descent approach to optimizing the network weights and it has a greater tendency than most other training systems to be trapped by local error minima without the ability to obtain an overall minimum error set of weights for the neural network. The robust backpropagation system, described by Riedmiller and Braun,[14] is a modified version of the gradient descent backpropagation training algorithm, has improved abilities to find a global optimum set of network weights and biases. The robust backpropagation technique uses only the signs of the partial derivatives of the error function with respect to each weight to determine the weight increment at the beginning of a training epoch. Levenberg-Marquardt training is based on Levenberg-Marquardt optimization techniques and utilizes the Jacobian matrix to determine corrective step increments at the start of each training epoch. Additional details have been given by Hagan and Menhaj.[15] This training technique converges rapidly to a global optimum set of weights and biases, but this system has a significantly larger memory requirement than most competing systems of training.[10,12-15]

A more rapid training technique is to form radial basis function networks (RBFNs). The hidden layer nodes of typical radial basis function networks are formulated as shown in Equation 4. Here N_j is node j, R is a radial basis transfer function shown in Equation 5, $\|x - w_j\|$ is the Euclidean vector norm of the difference between the vector of input nodes x and a weight vector w_j, and b_j is the bias for the node. The radial basis transfer function typically is a Gaussian function. Where typical transfer functions vary

from -1 to 1 over a range of $-\infty$ to $+\infty$, the radial basis transfer function varies from 0 to 1 over the same range with a maximum at zero.[12,16]

$$N_j = R(\|x - w_j\| \times b_j) \tag{4}$$

$$R(x) = e^{-x^2} \tag{5}$$

During training, data vectors are presented to the network and the output is compared to the correct value. When a new training set vector causes the network output to vary from the present margin of error, a new hidden layer node is established. Its weights and biases are mathematically determined to bring the network output into compliance. This process may yield an unusually high number of hidden layer nodes. If the number of hidden layer nodes approaches the number of training set vectors then the network is "memorizing" the training set and it is less useful for classification or multivariate interpolation.[12,16]

The k-nearest neighbor classifier assigns data vectors to categories based on the simple geometric assumption that data vectors of spectra or samples in a given category will resemble their class-mates to a greater degree than the other data vectors in the set. In turn, when the data vectors are treated as points in a geometric space, with the number of measurement dimensions equal to the dimensionality of the vectors. Various distance measurements may be used to gauge distance between data vectors (points) including Euclidean distance, Mahalanobis distance, and others. To assign the classification of a data vector, a distance matrix is formed, containing the distances between all pairs of data vectors in the data set. A data vector is assigned by the algorithm to belong to the class represented by the majority of a set of data vectors (numbering *k*) which are nearest in terms of distance to the vector being classified.[8,9,17]

The infrared spectra used in this study were assembled from a variety of sources, including the US Army, government contractors, and commercial companies. A total of 115 spectra were selected for examination in this study: 48 of organophosphorus pesticides and 67 of banned neurotoxins, precursors, and hydrolysis products. The original spectra were provided in a variety of formats including various JCAMP formats,[18] a proprietary US-Army format, and simple X,Y data pairs of wavenumber, absorbance

values. All were translated into a common format. It was necessary to restrict the spectral data used to a common frequency range, to adjust the normalization among the spectra, and to standardize the format. All spectra were restricted to the frequency range of the most restrictive set of spectra, resulting in a frequency of 650 – 2500 cm^{-1}.

The category number was assigned to reflect the true classification of the spectra, or was assigned to zero if the spectrum was being treated as a member of the test/evaluation set. Spectral variations due to the variety of instrumental and experimental measurement conditions were minimized by re-normalizing the spectral absorbance range prior to transduction.

The spectra were transduced into data vectors by dividing the spectral range into bins of equal frequency widths. The width of the spectral bins depended upon the number used to cover the frequency range. As an example, in the case where 200 bins were used, each bin had a width of 9.25 cm^{-1}. Other transducing variations were used with 100, 50, and 25 spectral bins spanning the frequency range.

Calculations

The transduced spectral data were preprocessed by calculating and applying Fisher weights as described by Sharaf, *et al.*[17] Data sets were visualized by performing principal component analyses[19] and plotting the first two principal components on a two-dimensional plot. Classifications were performed using a k-nearest neighbor algorithm (KNN)[17]and with artificial neural networks ANNs.[10-12,16] All calculations were carried out using MATLAB version 5.2.1 augmented with the Neural Network Toolbox version 3.0 and the Statistics Toolbox version 2.1.1.[20] Locally developed m-files were used to control calculations. An evaluation set was developed to test the stability and performance of the ANNs by adding synthetic noise to the original spectra. The synthetic noise was generated from a model developed by Schuchardt that predicted the infrared spectral noise to originate as a combination of Johnson noise, flicker noise, and shot noise, with the bulk of the noise arising from the first two sources.[22] Each ANN was trained using the training set data and was tested using both training and evaluation set data.

Results

Features were selected based either on Fisher weights or by eliminating selected spectral regions. Fisher weights were also used in some cases without any feature selection to weight the raw data features. When Fisher weights were used for selection, the weights were calculated for the features of the data set, and those features whose Fisher weights exceeded a given threshold value were retained for additional treatments.

Classifications were made using the KNN classification scheme. Distance matrices to support this classifier were computed using the Euclidean distance formula, which was available as a MATLAB function. This system provided a nonparametric classifier that was insensitive to the data set structure and could be applied to the data as a training set. The KNN classifier provided a number of misclassified objects, which indicate how separable the categories are for a given set of preprocessing and selection parameters.

In order to examine the effect of the bin size and number, the spectra were normalized to a constant magnitude and were binned into four data sets as described above. Fisher weights were calculated and applied to the raw feature bins. The resulting Fisher weights are plotted against frequency for these four data sets in Figure 3. The data sets were visualized by generating a joint principal components model and then plotting the first two principal components against each other. These plots are shown in Figure 4. The number of misclassified training and test set spectra, and the misclassification frequencies for the four binning sets are listed in Table I. In all cases the classifications were based on the 5-nearest neighbors of the data vector being assigned.

Neural Network Classification

Radial basis function networks

The four test/evaluation set pairs were evaluated with radial basis function networks. Each training set was used to generate and train a network. Training set objects belonging to the class of banned neurotoxins, precursors, and hydrolysis products were indicated by an output layer pattern of [1 0] and objects representing pesticides were indicated by an output layer of [0 1]. Outputs from the radial basis function networks (RBFNs) were real numbers, rounded to the nearest integers. A RBFN was produced

and trained for each of the four training. In each case a RBFN was obtained which classified the training set without error. The evaluation sets, based on noise-degraded spectra, were then entered into the trained RBFN. One spectrum from the 200-bin evaluation set and two spectra from the 25-bin evaluation set were misclassified. No errors occurred for the 100-bin and 50-bin evaluation sets.

Feed-forward Neural Networks

Feed-forward ANNs were produced by using the four training/evaluation sets. Tansigmoid transfer functions were used to carry input layer information to a 5-node hidden layer, and tansigmoid transfer functions carried information from the hidden layer to a 2-node output layer, which was related to the final output nodes by linear transfer functions. Weights for the layers were determined by various training functions following random initializations. The feed-forward ANNs were trained with a mean-square error goal of 1×10^{-5} and were allowed 300 training epochs to reach this goal. Three training techniques were used to train these networks, classical backpropagation, robust backpropagation, and Levenberg-Marquardt training. The training goal was not achieved within the allowed period by all three of the systems used. Results from the training process were not stable due to the random initialization of the network weights. The number of misclassified spectral patterns for the three techniques are given in Table I.

Levenberg-Marquardt training could only be applied to the training and evaluation sets obtained using 25 spectral bins. The training set was classified correctly in each of the five trials but 17 classification errors were produced from the evaluation set. Levenberg-Marquardt training could not be applied to networks using more input nodes due to memory limitations using a personal computer equipped with a 400 MHz Intel Pentium-II® microprocessor and 64 Mbytes of RAM.

All trials using gradient descent backpropagation training of the ANNs required all 300 training epochs and terminated with mean square errors greater than 1×10^{-2}. Errors in classification were noted from both the training and evaluation sets from these partially trained networks, summarized in Table I.

The memory requirements of robust backpropagation training were modest enough to allow it to be applied to all four data sets. Networks based on the 200-bin data set were trained successfully to mean

square error values less than the goal. In five trials, neural networks based on the 100-bin data set converged to the error goal in four cases, and narrowly missed it in the fifth trial, with a final error value of 1.16×10^{-5}. Three of five trial neural networks trained from the 50-bin data set failed to achieve the error goal, with the largest mean square error value being 1.04×10^{-4}. The largest error values were obtained from five networks based on the 25-bin input data, all of which failed to achieve the error goal, with the largest mean square error value being 1.14×10^{-3}. None of the neural networks produced in these trials exhibited any training set misclassifications. Table I summarizes the evaluation set misclassifications.

Discussion

The classification results obtained by the *k*-nearest neighbor classifier show little change among the data sets transduced at various resolutions between 9.25 cm^{-1} and 74 cm^{-1}. This result is consistent with Griffith [23] who showed that gaseous compounds can be measured and identified by FT/IR at resolutions as low as 50 cm^{-1}. The neural network classifiers are somewhat more sensitive to the resolution, and showed optimum classification from the 100-bin data set, with accuracies falling off as the transducing resolution was changed both positively and negatively from the optimum. Although neural network classifiers from the 25-bin data produced the most errors, the worst error frequencies from this data set were below 10 percent.

The radial basis function network classifier produced the best classification results. Of the neural network classifiers, the radial basis function networks were also the only stable classification result which could be reproduced by a repeated training from initial conditions. The radial basis function network could also be trained rapidly.

The neural network classifiers were all initialized with random network weights and they yielded trained networks which were somewhat unstable, since repeated training of the networks from initial conditions yielded differing results with each trial. The networks produced with simple backpropagation training yielded the greatest number of errors, and they were the only feed-forward neural network classifiers that produced erroneous classifications from the training set. The networks trained with robust

backpropagation converged to smaller mean-square-errors than those trained with simple backpropagation, and they produced no erroneous classifications from training set data. The evaluation set classifications produced by neural networks trained with the robust backpropagation system were slightly better than those produced by the simple backpropagation training. The 25-bin data set could be classified with all three feed-forward neural network systems tested. The best results from this data set were obtained when the neural network was trained with the Levenberg-Marquardt system, followed by the results from the networks trained by robust backpropagation and finally the networks trained with simple backpropagation. In view of the fact that the errors noted in Table I were due to multiple trials, the classifications obtained from most of the neural network classifiers were better than those obtained by the *k*-nearest neighbor classifier. The error frequency obtained by classifying the 25-bin data set with backpropagation training is slightly higher than those obtained for all of the data sets by the *k*-nearest neighbor classifier, but not significantly so.

In conclusion some information is lost as the spectral information is transduced at lower resolution, but sufficient information is retained to support classifications by *k*-nearest neighbor and neural network techniques even when the spectra are transduced at low resolutions. This is in general agreement with the findings of Griffith[23]. The *k*-nearest neighbor classifier is a classical pattern recognition technique which provides a performance measure with which to compare newer neural network classifiers. The *k*-nearest neighbor classifier is not particularly sensitive to the resolution of the data, with consistent results over a variety of bin size and resolution values. The neural network classifiers investigated here performed as well as the k-nearest neighbor classifier in nearly all cases. The feed-forward neural network classifiers were more sensitive to the data binning and resolution, with the optimum resolution differing with the network training technique. With the most effective training technique, even data with the lowest resolution provided successfully trained feed-forward networks able to classify the evaluation set with better than a 95% accuracy rate. Radial basis functions gave still better classification accuracies, but the networks reported herein showed signs of memorizing the training set data rather than forming an efficient decision surface within the data space. Additional fine-tuning of the radial basis function classifier is needed to reduce the tendency to memorize the data set. The neural network classifiers appear robust

enough to be used at lower data resolutions. The classification results obtained with the full spectral data width indicate that neural networks offer a promising means for classifying organophosphorus compounds. These classifications can also be performed quite accurately with radial basis function networks, and these may even be preferable if strict control of the network architecture and the number of nodes is not critical. Additional evaluations are underway to further reduce the number of features used in these classifications by selecting optimum frequencies. Feature selection results will be discussed elsewhere.

Acknowledgements

This work was supported as an in-house effort by Air Force Institute of Technology, Department of Engineering Physics and by Air Force Research Laboratory, Airbase and Environmental Technology Division. This work was funded by the US Army, ERDEC. We also thank US Army ERDEC and Bio-Rad Laboratories, Inc., Sadtler Division for providing infrared spectra used in this work. We are also grateful to other organizations and workers who provided infrared spectra or technical assistance but asked to remain anonymous.

References:

1. R. E. Porte, *Structure-Activity Predictions of Properties of Organophosphorus Pesticides and Chemical Agents Important for Optical Detection*, MS Thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, 1995.
2. CONVENTION ON THE PROHIBITION OF THE DEVELOPMENT, PRODUCTION, STOCKPILING AND USE OF CHEMICAL WEAPONS AND ON THEIR DESTRUCTION, (Corrected version in accordance with Depositary Notification C.N.246.1994.TREATIES-5 and the corresponding Procès-Verbal of Rectification of the Original of the Convention, issued on 8 August 1994).
3. M. T. Söderström and R. A. Ketola, *Fresenius Journal of Analytical Chemistry*, **350**, 162 (1994).
4. M. T. Söderström, R. A. Ketola, and O. Kostianen, *Fresenius Journal of Analytical Chemistry*, **352**, 550 (1995).
5. H. H. Willard, L. L. Merritt, Jr., J. A. Dean, and F. A. Settle, Jr., *Instrumental Methods of Analysis*, 7th Ed., Wadsworth Publishing Company, Belmont, CA, 1988.
6. P. R. Griffiths and J. A. de Haseth, *Fourier Transform Infrared Spectrometry*, John Wiley & Sons, New York, NY, 1986.
7. P. C. Jurs and T. L. Isenhour; *Chemical Applications of Pattern Recognition*, John Wiley and Sons, New York, NY, 1975.
8. B. R. Kowalski and C. F. Bender, *Journal of the American Chemical Society*, **94**, 5632 (1972).
9. B. R. Kowalski and C. F. Bender, *Journal of the American Chemical Society*, **95**, 686 (1973).
10. S. Haykin, *Neural Networks. A Comprehensive Foundation*, Macmillan College Publishing Company, New York, NY, 1994.
11. J. R. Long, H. T. Mayfield, M. V. Henley, and P. R. Kromann, *Analytical Chemistry*, **63**, 1256 (1991).
12. H. Demuth and M. Beale, *Neural Network Toolbox for Use with MATLAB. User's Guide Version 3*, The Mathworks, Inc., Natick, MA, 1998.
13. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Nature*, **323**, 533 (1986).

14. M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: The RPROP algorithm," *Proceedings of the IEEE International Conference on Neural Networks*, 586 (1993).

15. M. T. Hagan and M. B. Menhaj, *IEEE Transactions on Neural Networks*, **5**, 989 (1994).

16. S. Chen, C. F. N. Cowan, and P. M. Grant, *IEEE Transactions on Neural Networks*, **2**, 302 (1991).

17. M. A. Sharaf, D. L. Illman, and B. R. Kowalski, *Chemometrics*, John Wiley & Sons, New York, NY, 1986.

18. R. S. McDonald and P. A. Wilks, Jr., *Applied Spectroscopy*, **42**, 151 (1988).

19. E. R. Malinowski and D. G. Howery, *Factor Analysis in Chemistry*, John Wiley & Sons, New York, NY, 1980.

20. MATLAB version 5.2.1, The Mathworks, Inc., Natick, MA, 1998.

21. G. D. Schuchardt, *Automated Infrared Detection of Organophosphorus Compounds in Multicomponent Solutions*, MS Thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, 1995.

22. P. R. Griffiths, "FT-IR Spectrometry at Low Resolution: How Low Can You Go?", *Proceedings of the 9th International Conference on Fourier Transform Spectroscopy, SPIE Proceedings Series*, Volume 2089, J. E. Bertie and H. Weiser editors,1994, p. 2.

Table I. Summary of error rates from k-nearest neighbor classifier and five training and classification trials using feed forward neural networks trained with classical backpropagation training, robust backpropagation training, and Levenberg-Marquardt training.

KNN Classification Results, Number of Trials=1	200-bin		100-bin		50-bin		25-bin	
5-nearest Neighbor Classifier	Test	Train	Test	Train	Test	Train	Test	Train
Number Misclassified	5	8	5	8	7	7	6	8
Misclassification Frequency	0.04	0.07	0.04	0.07	0.06	0.06	0.05	0.07
Feed Forward Neural Network Results, Number of Trials=5								
Back-propagation of Errors								
Total Number Misclassified	12	11	20	14	19	22	43	40
Average Misclassification Frequency	0.021	0.019	0.035	0.024	0.033	0.038	0.075	0.07
Robust Back-propagation								
Total Number Misclassified	8	0	3	0	12	0	30	0
Average Misclassification Frequency	0.014	0	0.005	0	0.021	0	0.052	0
Levenberg-Marquardt Training								
Total Number Misclassified							17	0
Average Misclassification Frequency							0.03	0

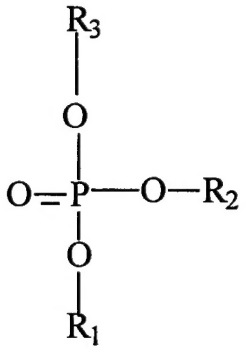
Figure Captions

Figure 1: Structures of common organophosphorus compound types, showing the type names. Reproduced from Porte [1].

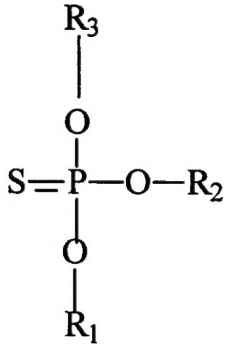
Figure 2: Structure and names of the four most common neurotoxins banned by the Chemical Warfare Convention.

Figure 3. Principal component factor score plots of the infrared spectral data transduced into varying numbers of bins over the frequency range of 650 – 2500 cm^{-1} with plot points labeled by class membership numbers. Class 1 members are banned neurotoxins, precursors, and hydrolysis products. Class 2 members are pesticides. (a) 200 bins. (b) 100 bins. (c) 50 bins. (d) 25 bins.

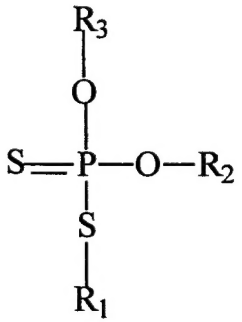
Figure 4. Fisher weights versus frequency (cm-1) for datasets obtained using four bin sizes: (a) 200 bins, 9.25 cm^{-1} wide. (b) 100 bins, 18.5 cm^{-1} wide. (c) 50 bins, 37 cm^{-1} wide. (d) 25 bins, 74 cm^{-1} wide.



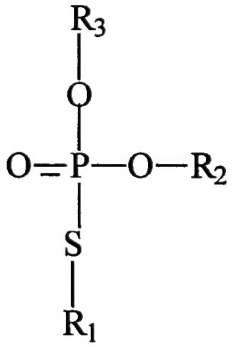
Organophosphate



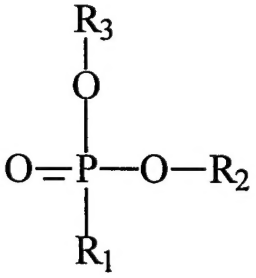
Organophosphorothionate



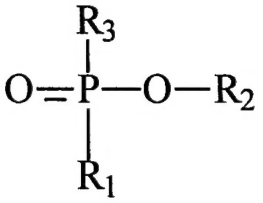
Organophosphorodithiate



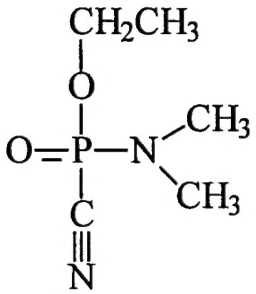
Organophosphorothiolate



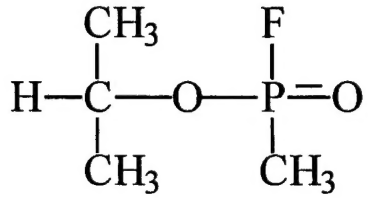
Organophosphonate



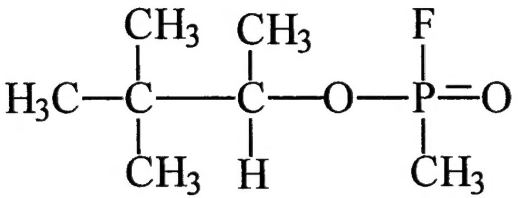
Organophosphinate



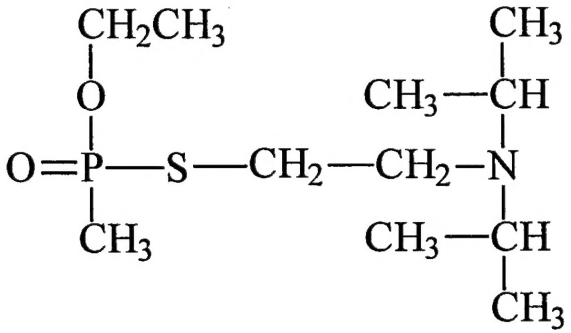
Dimethylphosphoramidocyanidic acid, ethyl ester



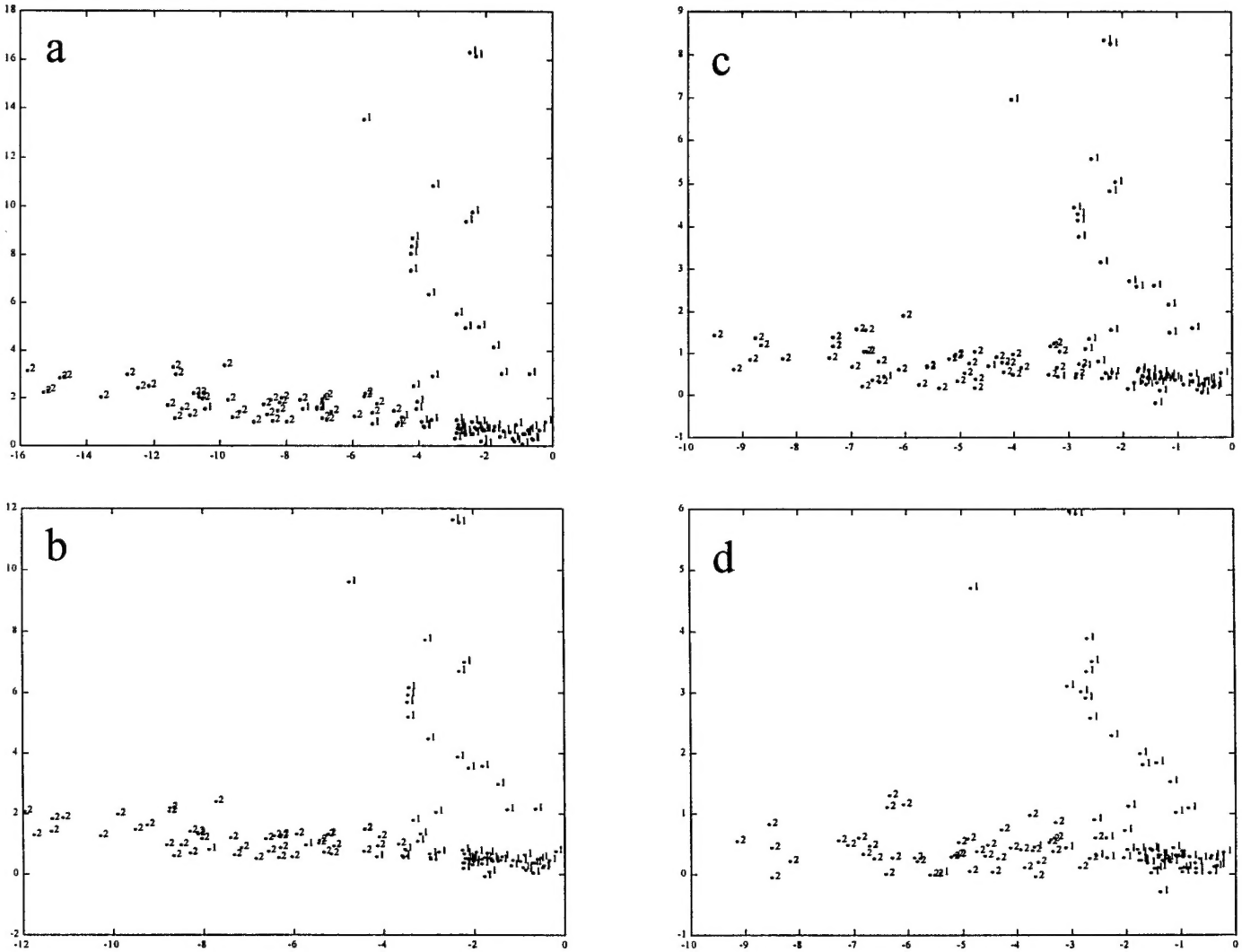
Methylphosphonofluoridic acid, (1-nethylethyl) ester

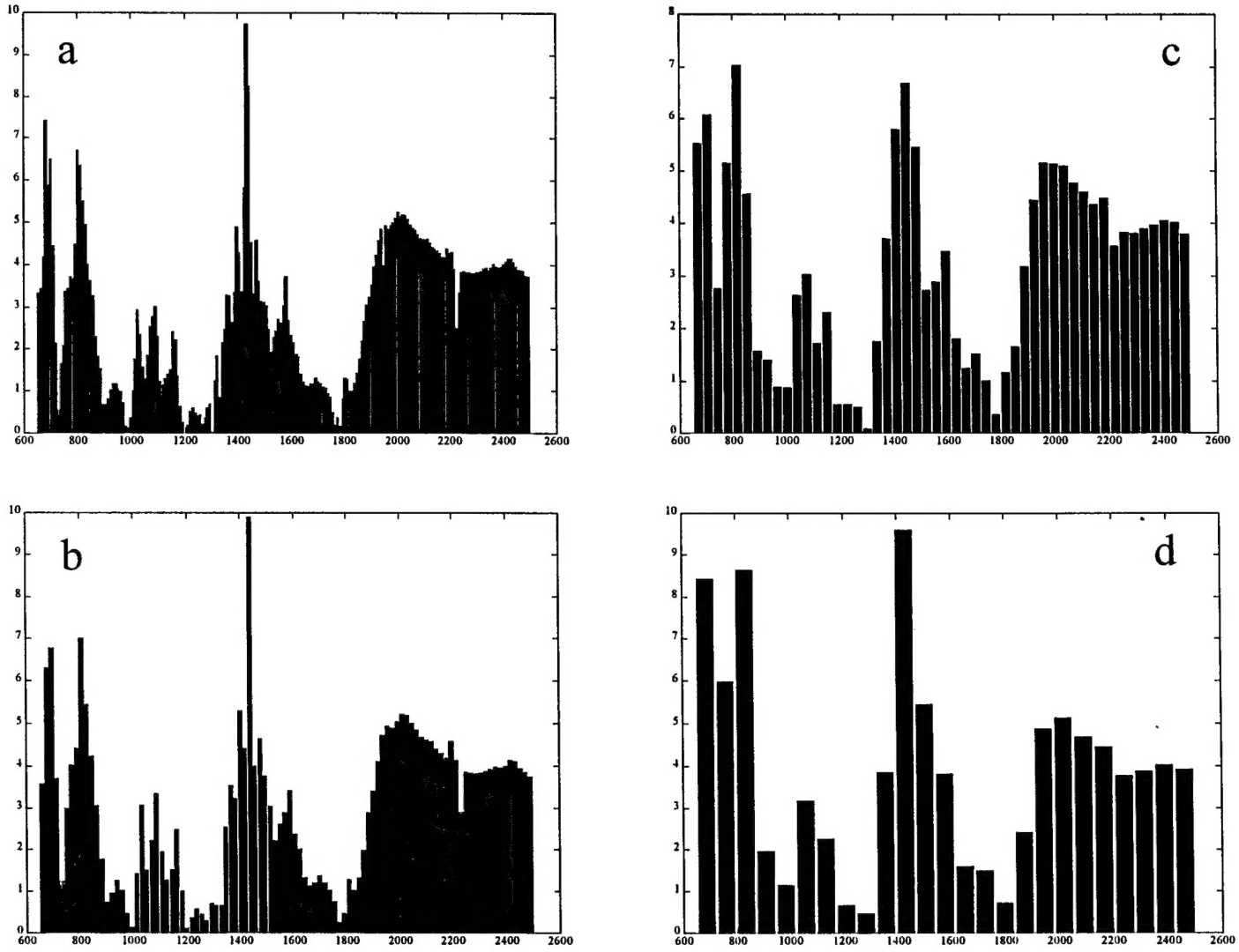


Methylphosphonofluoridic acid, 1,2,2-trimethylpropyl ester



Methylphosphonothioic acid, S-[2-[bis(1-methylethyl)amino]ethyl] O-ethyl ester


a
b
c
d


a
b
c
d
0
1
2
3
4
5
6
7
8
9
10
600
800
1000
1200
1400
1600
1800
2000
2200
2400
2600